



A comparison of recording modalities of P300 event-related potentials (ERP) for brain-computer interface (BCI) paradigm

Louis Mayaud, Marco Congedo, Aurélien van Laghenhove, Marjorie Figère,
Eric Azabou, F. Cheliout-Heraut

► To cite this version:

Louis Mayaud, Marco Congedo, Aurélien van Laghenhove, Marjorie Figère, Eric Azabou, et al.. A comparison of recording modalities of P300 event-related potentials (ERP) for brain-computer interface (BCI) paradigm. *Neurophysiologie Clinique = Clinical Neurophysiology*, 2013, 43 (4), pp.217-227. 10.1016/j.neucli.2013.06.002 . hal-00909021

HAL Id: hal-00909021

<https://hal.science/hal-00909021>

Submitted on 25 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A COMPARISON OF RECORDING MODALITIES OF P300 EVENT RELATED POTENTIALS (ERP) FOR BRAIN COMPUTER INTERFACE (BCI) PARADIGM

Authors:

Louis Mayaud (1-4), Marco Congedo (5), Aurélien Van Laghenhove (1,3), Marjorie Figère (1), Eric Azabou (2,3), Fawzia Cheliout-Heraut (3)

Affiliations:

- (1) INSERM, Centre d'Investigation Clinique et d'Innovation technologique (CICIT), UMR805, Garches, France
- (2) EA4497, Université Versailles Saint-Quentin (UVSQ), Versailles, France
- (3) Hôpital Raymond Poincaré, APHP, UVSQ, Garches, France
- (4) Institute of Biomedical Engineering (IBME), University of Oxford, Oxford, UK
- (5) GIPSA-Lab, CNRS, Grenoble University, Grenoble, France

Corresponding author:

Louis Mayaud
CICIT, Hôpital Raymond Poincaré,
104, bd Boulevard Raymond Poincaré
92380 Garches, France
Tel : +33650663491
Fax : +33147104633
Email : louis.mayaud@gmail.com

Running title: Comparing modalities of ERP extraction for BCI

Conflicts of interest: The authors have not transmitted any conflicts of interest.

Summary

Aims of the study A brain-Computer Interface aims at restoring communication and control in severely disabled people by identification and classification of EEG features such as Event Related Potentials (ERPs). The aim of this study is to compare different EEG recording modalities for extraction of ERPs. The first comparison evaluates the performance of six disc electrodes with that of the Emotiv headset, while the second evaluates three different electrode types (disc, needle, and large squared electrode).

Material and methods Ten healthy volunteers gave informed consent and were randomized to try the traditional EEG system (6 disc electrodes with gel and skin preparation) or the Emotiv headset first. Together with the six disc electrodes, a needle and a square electrode of larger surface were simultaneously recording near lead Cz. Each modality was evaluated over three sessions of auditory P300 separated by one hour.

Results No statically significant effect was found for the electrode type, nor was the interaction between electrode type and session number. There was no statistically significant difference of performance between the Emotiv and the six traditional EEG disc electrodes, although there was a trend showing worse performance of the Emotive headset. However, the Modality-Session interaction was highly significant ($p < 0.001$) showing that, while the performance of the 6 disc electrodes stay constant over sessions, the performance of the Emotiv headset drops dramatically between 2 and 3 hours of use. Finally, the evaluation of comfort by participants revealed an increasing discomfort with the Emotiv headset starting with the second hour of use.

Conclusion Our study does not recommend the use of one modality over another based on performance but suggests the choice should be made on more practical considerations such as the expected length of use, the availability of skilled labor for system setup and above all, the patient experience.

Keywords: P300, Event Related Potentials, EEG headset, Brain Computer Interface, Needle Electrode

Résumé

Objectifs de l'étude Les interfaces Cerveau-Ordinateur ont pour vocation la restauration de la communication et du contrôle chez les patients lourdement handicapés, via l'utilisation de caractéristiques EEG telles que les Potentiels Evoqués (PE). L'objectif de cette étude est la comparaison de différentes modalités d'enregistrement du signal EEG pour l'extraction des PE. Une première comparaison évalue la performance de six électrodes cupules avec celle d'un casque Emotiv, tandis qu'une seconde compare trois différents types d'électrodes (cupules, aiguilles et carrés de surface plus large).

Matériels et méthodes Dix volontaires sains ont donné leur consentement éclairé avant d'être randomisés pour essayer en premier, soit le système EEG traditionnel (six électrodes cupules avec gel et préparation de la peau), soit le casque Emotiv. Avec les six électrodes cupules, une aiguille et une électrode carrée de surface plus importante ont été utilisées près de Cz pour enregistrer simultanément l'activité EEG. Chaque modalité a été évaluée lors de trois sessions séparées d'une heure.

Résultats Aucune différence statistiquement significative n'a été trouvée, ni pour le type d'électrode, ni pour l'interaction entre le type d'électrode et la session. Il n'y a pas de différence significative entre le casque Emotiv et les six électrodes traditionnelles bien que la tendance montre de moindres performances pour le casque. L'interaction Modalité-Session s'est cependant révélée fortement significative ($p < 0.001$) ce qui s'explique par une baisse des performances du casque Emotiv entre la seconde et troisième heure d'utilisation. Enfin, l'évaluation du confort par les participants révèle une gêne importante associée à l'utilisation du casque Emotiv à partir de la seconde heure d'utilisation.

Conclusion Notre étude ne recommande pas l'utilisation d'une technique plutôt qu'une autre sur la base seule des performances mais suggère plutôt que ce choix soit dirigé par des considérations plus pratiques telles que durée attendue d'utilisation, disponibilité de personnel qualifié pour l'installation du système et surtout, le confort du patient.

Mots-clés: P300, Potentiels evoqués, casque d'électroencéphalographie, Interface Cerveau Ordinateur, Electrode aiguille

Introduction

A Brain-computer interface (BCI) is a direct communication pathway between the brain and an external device [1, 2]. The idea is challenging and involves sending discrete command to a device solely relying on the on-line analysis of brain signal, which completely bypasses the usual muscular and peripheral nerves output. Although first attempts concerned healthy users [3, 4] the research in BCI field has flourished in the 90s promoted by its potential use in restoring communication pathway for patients with severe motor disabilities, like spinal cord injuries or amyotrophic lateral sclerosis [5]. The quality of the EEG signal is of paramount importance to achieve a good BCI performance. Unfortunately, traditional gel-based sensors setup requires skilled technicians, is time consuming, and users' hair and electrodes have to be cleaned after use. It is therefore well recognized that one important improvement for the usability of BCIs would be the use of an easy-to-setup and no-gel-based head-mounted device.

Recently, several EEG recording devices with such characteristics have been proposed in the market. Because of low consumer price, wireless connection and the relatively high number of electrodes (14) the Emotiv headset (Emotiv, EPOC, Hong-Kong, HK) is particularly interesting [6]. It has been designed for gaming applications and therefore offers several advantages over traditional gel-based systems. In particular, electrodes are attached on a semi-rigid support, as seen on Figure 1, so that electrode positioning, while less accurate, is much faster. Emotiv headset has become increasingly popular in BCI community [7-10] and has already successfully been involved in several BCI designs [9, 11, 12]. Actual comparisons of the Emotiv Headset to more traditional EEG amplifier have already been reported for EEG recording [13] and Event Related Potentials (ERPs) extractions [14] indicating worse performance of the Emotiv Headset.

These studies were however not quantitative and a formal comparisons is necessary. In particular, it is unclear how quickly the signal quality drops while the felt pads are drying.

It has been showed that P300 ERPs can be used to select items displayed on a computer monitor with and may be able to provide a new communication channel to patients suffering from severe neurological or muscular disorders [15, 16]. These authors used a visual oddball paradigm, consisting in a 6x6 matrix where rows and columns are flashed at random. A complete repetition of flashes consists in all columns and rows flashing once (12 flashes total). Any of the 36 cells constituted potential targets, and is therefore flashed twice, once in its column and once in its row. These two rare events elicit a P300 response that can be identified with adequate algorithms. This paradigm was successfully tested with patients suffering from Amyotrophic lateral sclerosis (ALS) [17, 18] showing a stable P300 response. These results have been replicated several times [19, 20].

The P300 auditory event-related (endogenous) potentials (ERPs) were first described by Sutton et al. [21]. This positive component peaking 300 ms or later is elicited with a simple discrimination task, named “oddball paradigm”. In this task, two stimuli (tones) are presented in a random series such that one of them occurring relatively infrequently is the “oddball”. The subject has to distinguish between the tones by noting the occurrence of the target (pressing a button, mentally counting...) and not responding to the standard. In most cases, the frontal (Fz), central (Cz) and parietal (Pz) midline electrodes referenced at mastoids or earlobes have been used to obtain P300 potentials. A total of 200 stimuli randomly presented are classically used. The ERP are then averaged to increase the signal-to-noise ratio. The P300 potential is physiologically defined as the largest positive peak after the N100-P200-N200 complex within a latency window defined by the

sensory modality: auditory=250-450ms and visual=300-600ms. The P300 component is characterized by its amplitude (size) and latency (timing). Amplitude (μV) is defined as the voltage difference between the component peak and the pre-stimulus baseline. Latency (ms) is defined as the time from stimulus onset to the point of maximum positive amplitude of the ERP waveform occurring after the earlier exogenous components (N100, P200, N200).

The aim of this study is to compare different modalities of EEG recording for extraction of P300 ERPs. In a first experiment, the Emotiv headset was compared to a standard EEG recording system. In a second experiment, three different types of electrodes were compared. Comparison considered ease of use, comfort, setup time and ability to discriminate P300 target from non-target responses.

Materials and Methods

Participants

Participants were adult healthy volunteers of both gender recruited among nurses and research staff. Exclusion criteria were cognitive, eyesight or hearing deficit, history of epilepsy, pregnancy and people under judicial protection or without social security. The protocol received approval from both the local ethical board (Comité de Protection des Personnes, Saint-Germain-en-Laye, France) and the French regulatory agency (ANSN) and was conducted in compliance with good clinical guidelines and Helsinki declaration [22]. All participants included in the study gave their written informed consent.

Study Design

This study design was a randomized crossed-over pilot study with half-a-day wash out period. The main objective of the study is to compare different modalities for the characterization and classification of ERPs. Table I describes the modalities compared and reports the comparison criteria used.

Subjects were randomized (single block) to try DISC (six traditional disc electrodes) or EMOTIV (Emotiv headset) first. As detailed in Figure 2, the evaluation lasted half-a-day with three sessions separated by exactly an hour. Two evaluations for the same participant could not happen over two consecutive half-days.

DISC and EMOTIV were compared over two different sessions while electrodes dE (disc), nE (needle) and sE (squared of larger surface) are compared during the same session. Each evaluation was organized as follow:

- 1 Electrodes are set up by an experienced EEG technician (5-20 min);
- 2 The task is explained to the participant (5 min);
- 3 (3 times) Evaluation of the modality (15 min) + 1 hour break;

An analog visual scale (AVS, from 0 for “very uncomfortable” to 10 for “very comfortable”) was used to estimate the comfort during modality setup (before S0) and at the end of each session (S1, S2 and S3). Installation time was also recorded from case opening to beginning of the recording for each modality.

EEG recording

For Experiment I, the first modality (DISC) consisted of six standard silver/silver chloride disc electrodes (diameter = 10mm, surface = 31.4 mm^2 , Pactronic, France) that were set up at standard 10/20 locations Fz, Cz, Pz, Oz, PO3 and PO7 [23-25]. An abrasive paste (Nu-Prep Gel, Weaver and Company, Aurora, CO, USA) was applied to the scalp prior electrode positioning with a conductive paste (Tensive, Parker laboratories, USA). The impedance between the reference and each electrode was kept below 5 kOhms. Electrodes were connected to a g.MOBIlab+ (G.Tec, Guger Technologies, Graz, Austria) sampling EEG signals at 256 samples per seconds via a 16 bit A/D converter. The potential of all electrodes was referenced to the right mastoid. The left mastoid was used as ground. The amplifier was connected to a desktop computer (CPU: 2.4GHz, 4Go DDRMA2) running the OpenViBE open-source platform [26] (release 0.12.0-svn3107). For the second modality (EMOTIV), the Emotiv headset (Figure 1) [6] was set up front-side-back to allow collection of signals located rather in the parietal and occipital areas instead of frontal: the 14 single-use felt pads were irrigated with a 7% saline solution and approximately located at the extended 10/20 locations Fp1, F7, CP5, T5, P3, PO3, Fp2, F8, T4, CP6, P4, T6 and PO4 . The Emotiv headset uses a Common Mode Sense (CMS) electrode at F4 location and a Driven Right Leg (DRL) electrode at F3 that can be related to the ground and reference in more traditional acquisition systems. Electrodes impedances were controlled visually with the Emotiv Control Panel so that all sensors show “green”. Signals are internally digitized at 2048Hz (16-bit) and subsequently low pass filtered (43Hz) and down sampled to 128Hz before transmission to the acquisition module. The headset was connected wirelessly to a laptop (HP ProBook; Microsoft Windows 7 Professional SP1 32 bits; CPU: IntelCore i5 Dual Core m430 at 2.27GHz, Memory: DDR3 2Go) and interfacing OpenViBE platform.

For Experiment II, the following recording modalities were compared: a standard silver/silver chloride disc (diameter = 10mm, surface = 31.4 mm^2 , Pactronic, France) located at Cz, a silver/silver chloride needle electrode and a large silver/silver chloride square electrode (20x15mm, surface = 300 mm^2 , Pactronic, France) located at one centimeter from Cz on the right and left, respectively. Signals were recorded together with the DISC modality in Experiment I.

Auditory stimuli generation

For all comparisons, each session consisted of 20 independent blocks of one hundred stimulations at 5Hz: target (0.5kHz 0.1s-long beeps) and non-target stimuli (1kHz 0.1s-long beeps). To keep concentration high, participants were asked to note down the number of target stimuli they identified, while in each block was randomly populated with 15 ± 2 target stimuli. In total, each session consisted of 1700 non-target and 300 target stimuli.

EEG pre-processing

Raw EEG signals were stored as GDF format [27] with the type and location of auditory stimulations. For further data analysis, all EEG data were band-pass filtered between 0.01 and 30 Hz by means of Butterworth 4th order filters.

Experiment I: Comparison of 6 disc electrodes and the Emotiv Headset

For each session, the data consisted of a matrix X of dimension $N_t \times N_e$, where N_t is the number of temporal samples and N_e is the number of electrodes. Stimulations were coded in two vectors, t_i and y_i , of length $N_{stim} = 2000$, respectively indicating the temporal location (in number of sample from session start) and the type of stimulus (0 for non-target and 1 for target).

Cross-validation

For each session, the N_{stim} stimulations were split into five independent and random groups of equal size ($N_{val} = 400$) to allow a 5-fold cross-validation procedure. The K-fold validation is a standard technique used to estimate the variability of a result [28]. At each fold, a fifth of the data (N_{val}) is left aside for model validation (the “validation” set) while the remaining ($N_{train} = 1600$) is used for model fitting (the “training set”). At the end of this procedure, every observation has been used once and only once to estimate the performance of model (described below). For each fold, the training data was used to estimate the weights w_i of a spatial filter and the coefficients β_i for a linear classifier, according to a procedure detailed below. These were subsequently applied to each test set in order to estimate the overall model performance on unseen data.

Spatial filter

A spatial filter derives a waveform by linearly weighting the potential at each electrode. Spatial filters are designed to enhance interesting features of the data while suppressing the rest [29]. A few spatial filters are typically enough to describe the features of interests, thus the use of spatial filtering dramatically reduces the dimensionality of EEG data. Furthermore, spatial filters can be designed to output non-correlated waveforms, thus the classifier is provided with non-redundant data. The xDAWN algorithm [30] is a spatial filter that optimizes spatial weights W to enhance the

Signal to Noise Ratio (SNR) for ERPs elicited during and oddball paradigm. The filters are sorted by SNR enhancement ratio. The spatial filters were computed on the training data and applied to the validation set. Only the first three components of the xDAWN matrix were used, resulting in three spatial filters. W was therefore always of dimension $3 \times N_s$, leading to a transformed signal $X_{sp} = W * X$ of dimension $N_t \times 3$.

Model fitting

After each stimulation (t_i), a 1s time window was extracted from X_{sp} , decimated to 32 Hz and concatenated to create a feature vector d_i of dimension $1 \times 3 * N_s$, where $N_s = 32$. The feature vectors for all simulations constitute the dataset D of dimension $N_{stim} \times 3 * N_s$ associated with the binary target vector y_i .

The training dataset (320 targets, 1280 non-targets) was then used to fit an elastic net generalized linear model with Bernoulli link function (logistic regression) [31, 32] resulting in a vector of coefficients, β , of dimension $1 \times 3 * N_s$ where $\beta = 0$ for non-selected features. The number of features included in each model was forced to be $k = 30$.

Number of repetitions

The most commonly used technique to improve SNR in ERP analysis is that of averaging; it consist in presenting the same type of stimulus N_{rep} times and average target and non-target responses over all presentations [33]. To study the relation between the number of repetitions and the classifier performance for each modality, averaged responses were generated for a number of repetition going from $N_{rep} = 1$ to 30. This was achieved by randomly selecting N_{rep} target and $5 * N_{rep}$ non-target responses from the 80 target and 320 non-target responses present in the validation set. These randomly selected signals were averaged to create an averaged target

response $d_i^{N_{rep}}, y_i = 1$, and an averaged non-target response $d_i^{N_{rep}}, y_i = 0$, both of dimension $1 \times 3 * N_s$. This bootstrapping procedure was repeated 1000 times leading to a balanced data set $D^{N_{rep}}$ of dimension $2000 \times 3 * N_s$. In total, 30 different datasets $D^{N_{rep}}$ were generated to estimate the performance at each N_{rep} .

Estimation of performance

The coefficients β from the logistic regression model were finally applied to each datasets $D^{N_{rep}}$.

The probability of the i^{th} averaged response to be a target response (i.e. average of responses following a non-occurrent sound) is estimated as follow:

$$\hat{y}_i = \pi \left(\beta_0 + \sum_{j=1}^{3*N_s} \beta_j * d_{ij} \right)$$

where $\pi(x) = \frac{1}{1+e^{-x}}$ is the logistic function with zero mean and unit variance.

The Area Under the Receiver Operating Curve (AUROC) was used to estimate the discriminative power of the model on the validation set. It can be seen as the probability of any i^{th} target response ($\forall i / y_i = 1$) to be given a higher probability of being a target response, \hat{y}_i , than any j^{th} non-target response ($\forall j / y_j = 0$) [34] and is noted $AUROC = P(\hat{y}_i > \hat{y}_j)$.

Experiment II: Comparison of disc (dE), needle (nE) and squared (sE) single-electrode

For each stimulus, a one-second time-window starting at the stimulation trigger was extracted for target and non-target stimuli for the three electrodes of interest (at Cz). ERPs were averaged within each session, resulting in two waveforms for each electrode, for each session and for each participant. Modalities were compared using the root mean square difference (RMSd) after signal normalization, computed such as

$$\text{RMSd}_e(t) = \sqrt{\frac{\sum_{t=1}^{N_s} (S_e^{\text{target}}(t) - S_e^{\text{Ntarget}}(t))^2}{N_s}}$$

where e is the electrode index, t is the sample number, N_s the number of samples in the response window, and $S_e^{\text{target}}(t)$ and $S_e^{\text{Ntarget}}(t)$ the n^{th} sample for electrode e for the target and non-target response, respectively.

Statistical analysis

For Experiment I we conducted the following tests:

- the AUROC was ranked and used as a dependent variable in a non-parametric repeated measure ANOVA using permutation methods [35] to account for data non-normality ; Modality (EMOTIV or DISC) and session number were used as crossed-over random factor; the 5-fold and the number of repetitions ($N_{rep} = \{5,10,15\}$) were defined as nested random factors (at each session);
- Difference in satisfaction at installation and after sessions 1, 2 and 3, were tested with an ANOVA with modality and session number as the two repeated-measure factors [35-37];
- Comparison of setup time for the two modalities was compared with a paired t-test.

For all statistical tests the tolerance for type I error was set to 0.05. All analyses were performed using Matlab (Version 8.0.0.783 - R2012b) and toolboxes referenced in this section.

For Experiment II, the RMS of the difference (RMSd) was used as a dependent variable in a repeated measure ANOVA with electrode type (disc, needle, or square) and session number as repeated-measure factor. The repeated measures ANOVA uses a multivariate framework (Hotelling T-square) that automatically account for correlation between measures [37] and therefore does not need to account for sphericity [36].

Results

Subject

Ten healthy volunteer aged 24 to 58 (mean 36.3 ± 12.2), five male (50%), were included in the study after giving informed consent.

Waveform analysis

Table II summarizes means latencies of the recorded P300 responses of each participant during each session as well as session's means. These data ranges are consistent with those existing in the literature [38]. Furthermore, analysis of P300 responses for each participant revealed a significant relation between P300 latencies and ages of the subjects ($p=0.02$) [39] that can be described with the following equation: $latency = 176.4 + 4.9 \times age$. Figure 3 (Top) plots the grand averages P300 responses recorded during each one of the three sessions for all the 10

studied subjects using each type of electrode. Figure 3 (Bottom) represented computed scalp topographic mapping of P300 responses peak amplitude. Table III shows average P300 peak amplitude data for the three electrode types used (dE, sE and nE) for all participants across the three different sessions.

Experiment I: Comparison of EMOTIV and DISC

Table IV presents the results from the non-parametric analysis of variance comparing the performance of the 6 disc electrodes (DISC) and the Emotiv Headset. (EMOTIV) The modality factor was not found to be statistically significant although there is a trend for the Emotiv headset to provide worse results as illustrated in Figure 4. The Session factor and Modality-Session interaction were significant, as seen in Table IV, and reflects the drop in performance for the Emotiv headset between sessions two and three that can be observed in Figure 4.

It took on average 26.3 ± 4.6 minutes to setup 6 standard disc electrodes against 11.5 ± 3.5 minutes for the Emotiv headset ($t(9)=7.5$; $p<0.001$). Figure 5 presents the evolution of comfort rated by participants at installation and after each session for the Emotiv headset and the set of 8 electrodes (including the square and the needle electrodes).

For comparison of comfort, modality, session and modality-session interaction were all found to be statistically significant factors as seen in Table V. This means that the Emotiv headset significantly is less comfortable than the disc electrodes and also that its comfort decreases with time.

Issues mentioned by participants for DISC included an unpleasant feeling during the needle electrode setup (5/10) and an itchy feeling from the nat (3/10). Eight complained of severe pain while using the Emotiv headset after an hour; they mentioned a “painful” temporal pressure (6/10), frontal (1/10) or a “severe radiating pain in the jaw” (1/10).

Experiment II: Comparison of modalities dE, nE and sE

Figure 6 shows the evolution of RMSd over three sessions for electrode type dE, nE and sE. It shows that all electrode types provide similar performance over the three sessions. The ANOVA show statistically significance for none of the factors nor their interaction.

Discussion

Waveform analysis

Amplitudes, latencies, waveforms as well as scalp topographic distribution of auditory P300 responses recorded in this study are largely in accordance with those reported in the literature [23-25]. Interestingly, the well-known correlation between age of subject and P300 latency was noticed in this study [39].

Experiment I: Comparison of EMOTIV and DISC

Results show that there is no overall statistically significant difference in ERP extraction between the six standard disc electrodes and the 14 wet electrodes from the Emotiv headset even though

there is a trend for the Emotiv to provide worse results, as reported in previous studies [12-14].

The fact that Modality is not significant can be explained by the large variability of results with the Emotiv headset that can be observed in Figure 4. The Modality-Session interaction factor was however found significant suggesting that the signal from the Emotiv headset degrades much faster over time. This phenomenon is illustrated on the right hand side of Figure 4, which shows poorer performance for the Emotiv headset (green) after two hours, together with a much wider variability amongst participants.

Fundamental differences in the design of the two recording modalities can explain the better performance of the traditional EEG system. First of all, Emotiv headset electronic specifications (noise or entry impedance) are unknown and are likely to underperform the one of the Gtech MOBILab amplifier that was designed for electrophysiological studies and not gaming and entertainment. Then, Emotiv raw EEG signals are transferred wirelessly for processing and storage, which possibly involves packet loss (even though the receptor was located as near as possible to the headset) and further software corrections. Electrode location can also account for lack of discriminatory power, since the current headset does not cover well the central-parietal area. This could however easily be corrected in future design of headset. Finally, the electrode type (wet felt pads with no skin preparation) likely accounts for the largest part of the difference. In particular, wet electrode drying is a plausible mechanism to explain the poor performance of the Emotiv headset in the final session.

On the other hand, the one-meter long non-shielded cable used to connect disc electrodes to the MOBILab amplifier constitutes an important source of electromagnetic noise, while the Emotiv headset digitizes signals nearly on-site. Also, the Emotiv headset benefits from a larger number of

electrodes (14 against 6), which favors good performance of the spatial filter. In particular, it has been shown that using a spatial filter eight disc electrodes suffice to obtain good performances [25], but this could be done at the cost of a longer setup time. The analysis was therefore re-run considering only the 6 best located electrodes available on the Emotiv headset (CP5, P3, PO3, C4, P4 and PO4) and showed a statistically significant effect of modality ($p(\text{MC}) < 0.05$). The difference in performance can still partly be explained by the different electrode locations though other aspects of the design reasonably account for a significant part of it.

Indeed, comparing two EEG recording modality aiming at their practical use involves other criteria than performance only. Setup time can be particularly important, favoring the Emotiv headset that is more than twice faster to setup (11 against 26 minutes, $p < 0.001$).

The use of the Emotiv headset was unfortunately associated with major discomfort in all but two participants (80%). All participants however ended the three sessions despite being said they could drop out at any time. We believe that the source of discomfort was from the two electrodes directly attached to the C-arm, on which most of the pressure was applied. Even though such headset offers advantages in terms of ergonomic and ease of use, the pain associated with its prolonged use makes it inappropriate for patients who cannot express pain, who are unfortunately some of the patients interested in BCI applications.

Experiment II: Comparison of modalities nE, dE and sE

The RMSd was used instead of the P300 peak amplitude to compare the electrode types for several reasons. Firstly, the discriminative information in ERP response is spread at different latencies [40, 41], which would not be captured by the amplitudes at a point in time. Second,

there are some caveats associated with extraction and comparison of ERPs amplitudes [42] including problems related to scaling [43] and sensitivity to noise, which are circumvented by using the RMSd approach.

Results suggest that all three traditional electrodes located at Cz provide equivalent performance over the three sessions, which confirms that needle electrode provide equivalent performance to disc electrodes [44]. Similarly, the squared electrode offering a larger surface of contact does not outperform the traditional disc electrode. We conclude that for P300 studies lasting up to two hours, there is no clear advantage in term of signal quality of one modality over another. The choice of recording modality is therefore left to practical consideration such as patient comfort and setup time. In particular, needle electrodes, despite a short installation time, should not be used in quadriplegic patients, potential BCIs users, who often develop scalp hyper-sensibility.

Conclusion

The aim of our study was to compare different modalities for extraction of ERPs during an auditory oddball paradigm. The first experiment compares the Emotiv headset with six gel-based disc electrodes: it shows no overall statistically significant difference between the two modalities although there is a trend for the Emotiv headset to provide worse performance. Our results also suggest that signals from the Emotiv headset degrade significantly between the second and third hour of use. On average, the setup time for the Emotiv headset was twice shorter. Finally, the Emotiv headset was associated with an important discomfort in eight participants after an hour. The second experiment compares three types of electrodes (discs, needle and squared with larger surface) showing no clear advantage of one technique over another. To conclude, results from

this study highlight benefits and pitfalls for each modality without clear winner: the choice of modality should be decided on a case-by-case basis given: expected time of use, the availability of skilled technicians for setup, desired performance and above all patient's comfort. In a clinical context however we discourage the use of the Emotiv headset for BCI as it potentially can be painful. We believe our results can guide the design of future EEG headsets.

Acknowledgment

We would like to thank the National Research Agency (ANR) and the General Directorate for Armament (DGA) for funding this project (Project ANR- 09-TECS-013-01-RoBIK). We would also like to thank Michelle Lejaille and Sandra Pottier for their support in setting up this study. Finally we particularly appreciated the expertise and support of Catherine, Gilles and Jean-Marie at the Physiology Department.

References

1. Allison, B.Z., E.W. Wolpaw, and J.R. Wolpaw, *Brain-computer interface systems: progress and prospects*. Expert review of medical devices, 2007. **4**(4): p. 463-474.
2. Wolpaw, J.R., et al., *Brain-computer interfaces for communication and control*. Clinical Neurophysiology, 2002. **113**(6): p. 767-791
3. Vidal, J.J., *Toward direct brain-computer communication*. Annual review of Biophysics and Bioengineering, 1973. **2**(1): p. 157-180
4. Vidal, J.J., *Real-time detection of brain events in EEG*. Proceedings of the IEEE, 1977. **65**(5): p. 633-641
5. Birbaumer, N. and L.G. Cohen, *Brain-computer interfaces: communication and restoration of movement in paralysis*. The Journal of physiology, 2007. **579**(3): p. 621-636
6. Emotiv, E., *Software Development Kit*, 2010.
7. Campbell, A., et al., *NeuroPhone: brain-mobile phone interface using a wireless EEG headset*2010: ACM. 3-8
8. Moore Jackson, M. and R. Mappus, *Applications for brain-computer interfaces*. Brain-Computer Interfaces, 2010: p. 89-103.
9. Stamps, K. and Y. Hamam, *Towards inexpensive BCI control for wheelchair navigation in the enabled environment—a hardware survey*. Brain Informatics, 2010: p. 336-345.
10. Zhang, B., J. Wang, and T. Fuhlbrigge, *A review of the commercial brain-computer interface technology from perspective of industrial robotics*2010: IEEE. 379-384
11. Thobbi, A., R. Kadam, and W. Sheng, *Achieving Remote Presence using a Humanoid Robot Controlled by a Non-Invasive BCI Device*. International Journal on Artificial Intelligence and Machine Learning. **10**: p. 41-45.
12. Liu, Y., et al. *Implementation of SSVEP based BCI with Emotiv EPOC*. in *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2012 IEEE International Conference on*. 2012. IEEE.
13. Stytsenko, K., E. Jablonskis, and C. Prahm. *Evaluation of consumer EEG device emotiv epoc*. in *MEi: CogSci Conference 2011, Ljubljana*. 2011.
14. Duvinage, M., et al. *A P300-based Quantitative Comparison between the Emotiv Epoc Headset and a Medical EEG Device*. in *Biomedical Engineering/765: Telehealth/766: Assistive Technologies*. 2012. ACTA Press.
15. Farwell, L.A. and E. Donchin, *Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials*. Electroencephalography and clinical neurophysiology, 1988. **70**(6): p. 510-523
16. Donchin, E., K.M. Spencer, and R. Wijesinghe, *The mental prosthesis: assessing the speed of a P300-based brain-computer interface*. IEEE Trans Rehabil Eng, 2000. **8**(2): p. 174-9.
17. Sellers, E.W. and E. Donchin, *A P300-based brain-computer interface: initial tests by ALS patients*. Clin Neurophysiol, 2006. **117**(3): p. 538-48.
18. Sellers, E.W., A. Kubler, and E. Donchin, *Brain-computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller*. IEEE Trans Neural Syst Rehabil Eng, 2006. **14**(2): p. 221-4.
19. Mak, J.N., et al., *EEG correlates of P300-based brain-computer interface (BCI) performance in people with amyotrophic lateral sclerosis*. Journal of neural engineering, 2012. **9**(2): p. 026014 %@ 1741-2552.

20. Nijboer, F., et al., *A P300-based brain–computer interface for people with amyotrophic lateral sclerosis*. Clinical Neurophysiology, 2008. **119**(8): p. 1909-1916
21. Sutton, S., et al., *Evoked-potential correlates of stimulus uncertainty*. Science, 1965. **150**(3700): p. 1187-8.
22. Rickham, P., *Human experimentation. Code of Ethics of the World Medical Association. Declaration of Helsinki*. British Medical Journal, 1964. **2**(5402): p. 177.
23. Duncan, C.C., et al., *Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400*. Clinical Neurophysiology, 2009. **120**(11): p. 1883-1908.
24. Klem, G.H., et al., *The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology*. Electroencephalography and clinical neurophysiology. Supplement, 1999. **52**: p. 3.
25. Cecotti, H., et al., *A robust sensor-selection method for P300 brain–computer interfaces*. Journal of neural engineering, 2011. **8**(1): p. 016001.
26. Renard, Y., et al., *OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments*. Presence: teleoperators and virtual environments, 2010. **19**(1): p. 35-53
27. Schlögl, A., *GDF-A general dataformat for biosignals*. arXiv preprint cs/0608052, 2006.
28. Lotte, F., et al., *A review of classification algorithms for EEG-based brain–computer interfaces*. Journal of neural engineering, 2007. **4**.
29. Congedo, M., C. Gouy-Pailler, and C. Jutten, *On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics*. Clinical Neurophysiology, 2008. **119**(12): p. 2677-2686
30. Rivet, B., et al., *xDAWN algorithm to enhance evoked potentials: application to brain–computer interface*. Biomedical Engineering, IEEE Transactions on, 2009. **56**(8): p. 2035-2043.
31. Friedman, J., T. Hastie, and R. Tibshirani, *glmnet: Lasso and elastic-net regularized generalized linear models*. R package version, 2009. **1**.
32. Park, M.Y. and T. Hastie, *L1- regularization path algorithm for generalized linear models*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2007. **69**(4): p. 659-677.
33. Shipton, H.W. *A simple averaging technique for the study of evoked cortical potentials in man*. in *Medical Electronics: Proceedings of the Second International Conference on Medical Electronics, Paris, 24-27 June 1959*. 1960. Iliffe.
34. B, H.J.a.M., *The meaning and use of the area under a receiver operating characteristics (ROC) curve*. Radiology, 1982. **143**(1).
35. Anderson, M.J., *A new method for non- parametric multivariate analysis of variance*. Austral Ecology, 2001. **26**(1): p. 32-46
36. Pernet, C.R., et al., *Limo EEG: a toolbox for hierarchical linear modeling of electroencephalographic data*. Computational intelligence and neuroscience, 2011. **2011**: p. 3
37. Rencher, X., *Methods of Multivariate Data Analysis*, 1995.
38. Delorme, A. and S. Makeig, *EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis*. Journal of neuroscience methods, 2004. **134**(1): p. 9-21.
39. Polich, J., L. Howard, and A. Starr, *Effects of age on the P300 component of the event-related potential from auditory stimuli: peak definition, variation, and measurement*. Journal of Gerontology, 1985. **40**(6): p. 721-726.

40. Aloise, F., et al., *A comparison of classification techniques for a gaze-independent P300-based brain-computer interface*. J Neural Eng, 2012. **9**(4): p. 045012.
41. Manyakov, N.V., et al., *Comparison of classification methods for P300 brain-computer interface on disabled subjects*. Comput Intell Neurosci, 2011. **2011**: p. 519868.
42. Todd, R.M., et al., *The time course of social-emotional processing in early childhood: ERP responses to facial affect and personal familiarity in a Go-Nogo task*. Neuropsychologia, 2008. **46**(2): p. 595-613.
43. Haig, A.R., E. Gordon, and S. Hook, *To scale or not to scale: McCarthy and Wood revisited*. Electroencephalography and clinical neurophysiology, 1997. **103**(2): p. 323-325.
44. Zablow, L. and E.S. Goldensohn, *A comparison between scalp and needle electrodes for the EEG*. Electroencephalography and clinical neurophysiology, 1969. **26**(5): p. 530-533.

Figures

Figure 1: The Emotiv EEG headset (with permission).



Figure 2: Description of the protocol. Each participant tries all recording modality (DISC, dE, nE, sE and EMOTIV) with a different order. Each modality is evaluated over half-a-day with three different sessions (S1, S2 and S3) separated by one hour.

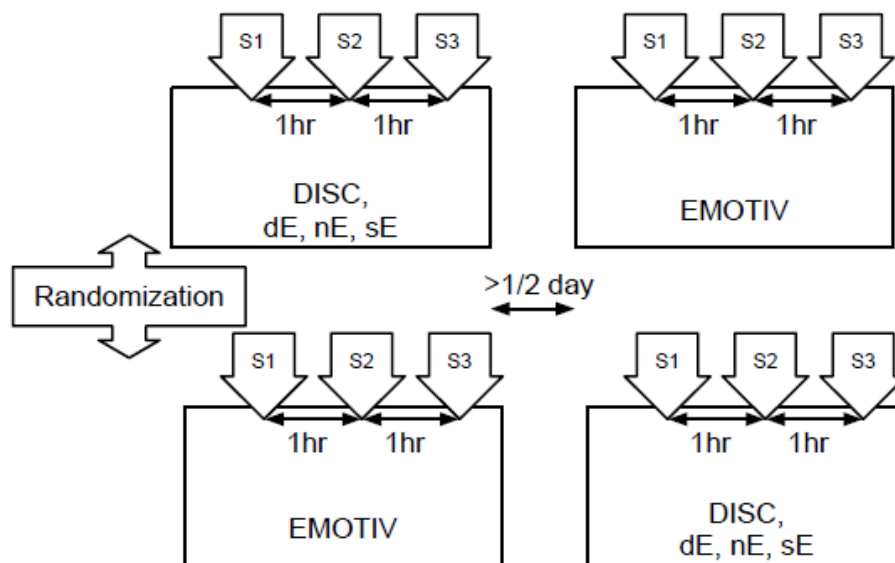


Figure 3: For each session: (Top) Grand average (N=10) P300 responses on disc electrode (dE in blue), square electrode (sE in green) and needle electrode (nE in red) with grand average (N=10)

non-target response (dE in dotted black); (Bottom) topographic map of grand average (N=10) amplitude at the P300 peak.

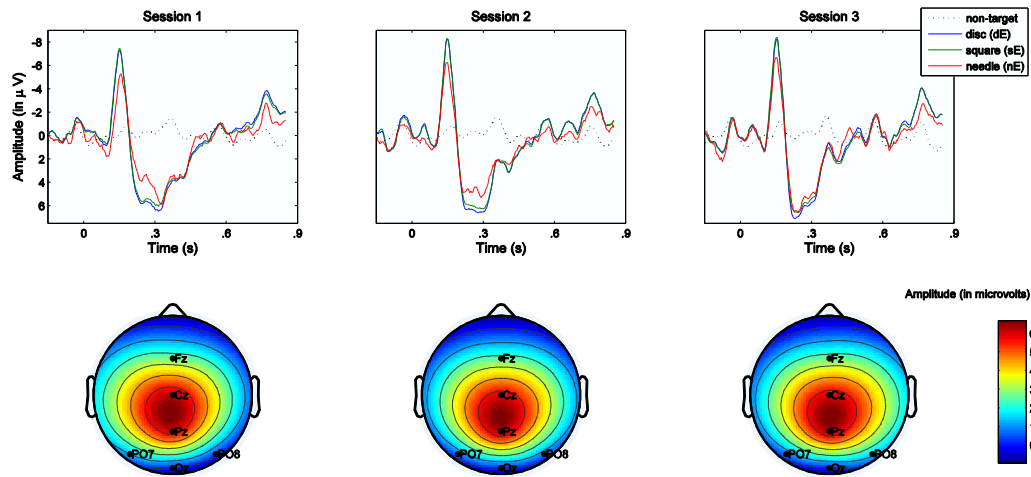


Figure 4: Evolution of the median Area Under the Receiver Operating Curve (AUROC) against number of repetitions of the target stimulus over three sessions for the Emotiv headset (grey) and six standard disc electrodes (black). Error bar show the standard deviation of the results.

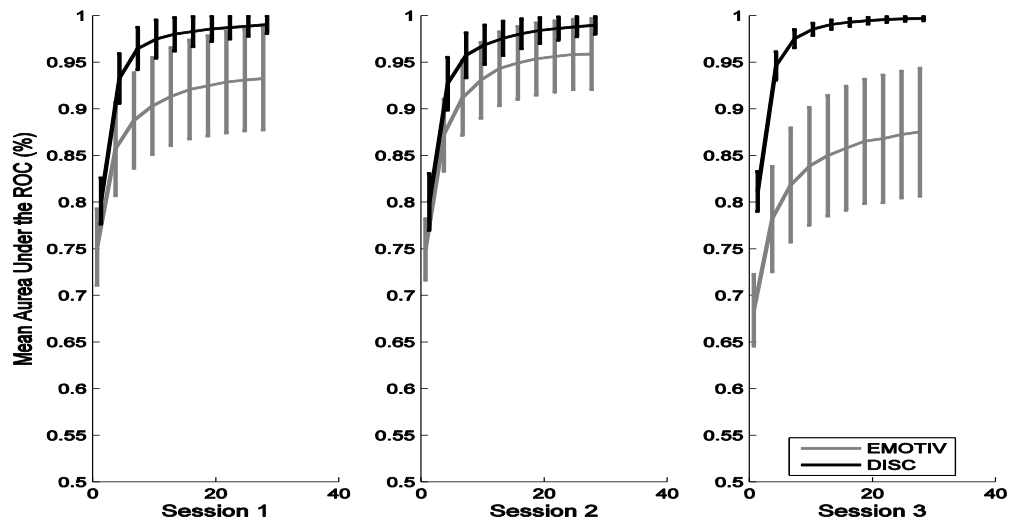


Figure 5: evaluation of comfort with Analogous Visual Scale (AVS) for 6 disc electrodes (DISC in blue) and the Emotiv headset (EMOTIV in green) at setup and after the first, second and third session. Error bars represent the standard deviation across all participants.

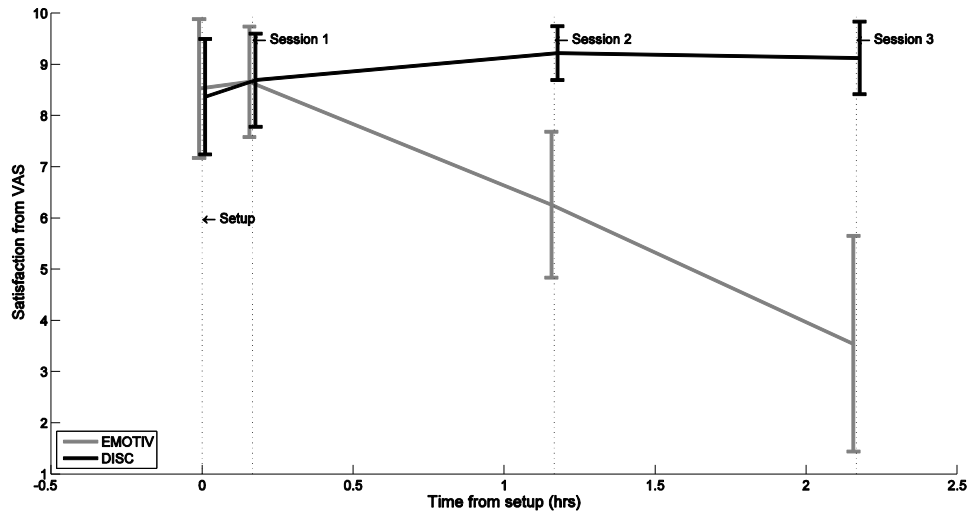
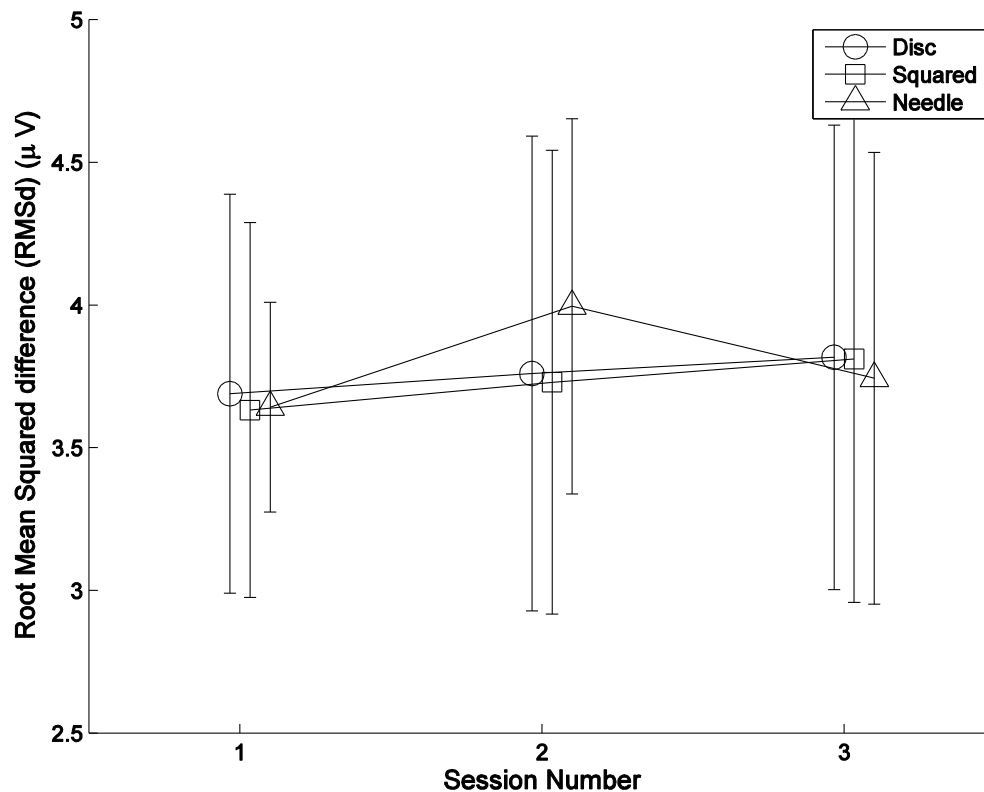


Figure 6: Evolution of the average RMSd over three sessions for discs (circle), squared electrode (square) and needle electrode (upright triangle) located at Cz. The error bar represents the standard deviation for the value across participants.



Tables

Table I: Description of the different recording modalities and criteria used for comparison

Recording modalities compared	Criteria for comparison
EXPERIMENT I	
<ul style="list-style-type: none"> • DISC: 6 traditional silver/silver chloride disc electrodes • EMOTIV: Emotiv Epoch headset (14 gold-plated electrodes in contact with a humid cotton-pad) 	<ul style="list-style-type: none"> • Ability to discriminate between target and non-target stimuli, the Area Under the Receiver Operating Curve (AUROC). • Ease of use and evaluation of comfort
EXPERIMENT II	
<ul style="list-style-type: none"> • dE : 1 standard silver/silver chloride disc electrode (Cz) • nE : 1 needle electrode (Cz) • sE : 1 large square silver/silver chloride electrode (Cz) 	<ul style="list-style-type: none"> • Root mean square difference (RMSd) between target and non-target responses.

Table II: P300 latency (ms) for all subjects across different sessions with individual and session means. Peaks were identified from waveforms by a senior neurophysiologist (FH).

	Gender	Age	Session 1	Session 2	Session 3	Individual mean
Participant 1	M	25	268,2	253,4	252,4	258.0±8.8
Participant 2	F	24	331,4	328,7	304,3	321.5±14.9
Participant 3	M	57	262,9	280,3	270,8	271.3±8.7
Participant 4	M	48	380,3	318,5	324,7	341.2±34.0
Participant 5	F	32	338,5	363,1	272	324.5±47.1
Participant 6	F	36	298,4	287,3	267,6	284.4±15.6
Participant 7	F	27	284,7	262,9	285,3	277.6±12.7
Participant 8	F	58	443,4	425,7	435,1	434.7±8.9
Participant 9	M	27	292,7	323,2	265,8	293.9±28.7
Participant 10	M	42	356,9	348,6	351,3	352.3±4.2
Session mean	50% (M)	37±13	325.7±56.5	319.2±51.9	302.9±55.5	315.9±53.7

Table III: Average P300 peak amplitude for three electrode types (dE, sE and nE) for all participants across different sessions.

	Session 1	Session 2	Session 3	Mean
Disc (dE)	6.4 ± 3.0	7.4 ± 2.7	7.5 ± 1.9	7.1 ± 0.6
Squared (sE)	6.1 ± 2.8	7.2 ± 2.6	7.1 ± 1.8	6.8 ± 0.6
Needle (nE)	5.9 ± 2.2	5.6 ± 2.7	6.6 ± 1.9	6.0 ± 0.4

Table IV: Results from the non-parametric analysis of variance comparing the AUROC (independent variable) of the 6 disc electrodes and the Emotiv headset (Mo) over three sessions (Se) for different folds (Kfold).

Source	df	SS	MS	F-statistic	P-value (permut.)	P-value (MC)
Modality (Mo)	1	1.56	1.568	10.3	0.11	0.09
Session (Se)	2	0.13	0.066	20.1	<0.001	<0.001
Kfold	24	0.08	0.003	0.5	0.98	0.98
(MoxSe)						
Number of Repetitions (MoxSexKfold)	60	0.42	0.007	0.5	1.00	1.00
MoxSe interaction	2	0.30	0.158	46.1	<0.001	<0.001
Residual	810	12.6	0.0155			
Total	899	15.1				

Abbreviations: df, degrees of freedom; SS, Sum square; MS, Mean Square; MC, Monte-Carlo

Table V: results from the repeated measures ANOVA for analysis of comfort.

Source	Df	T ² statistic	F statistic	P-value
--------	----	--------------------------	-------------	---------

Modality	1	23.2	23.2	<0.01
Session	3	45.6	12.4	0.01
Modality x Session	3	9319.2	187.0	<0.01

Abbreviations: df, degrees of freedom.